

Analysis of Autonomic Regulation in Cancer Survivors During Daily Physical Activity: A Real-World Wearable ECG Study

Sajad Farrokhi¹, Lerick Sequeira², Shanna L. Burke², Waltenegus Dargie³, *Senior Member, IEEE*, and Christian Poellabauer⁴, *Senior Member, IEEE*

Abstract—This study investigates heart rate (HR) and heart rate variability (HRV) responses to physical activity in breast cancer survivors using wearable electrocardiogram (ECG) data collected in real-world settings. Reliable HRV analysis in such environments is challenging due to motion artifacts and activity-related signal degradation. To address this, we use an approach that combines accelerometer and gyroscope data for activity intensity segmentation (light, moderate, vigorous) with a robust ECG processing pipeline incorporating R-peak detection and annotation-free signal quality assessment. Because vigorous activity produced unreliable HRV estimates, analyses focused on light and moderate activity levels. Using 30 s, 1 min, and 2 min windows, HR and HRV metrics were computed and compared between breast cancer survivors and healthy controls. Compared with healthy controls, cancer survivors generally exhibited higher HR and lower HRV across activity levels. During light activity, HR increased from 95.7 bpm in controls to 103.4 bpm in cancer survivors. Differences became more pronounced during moderate activity, where RMSSD decreased from 39.7 ms to 22.1 ms and SDRR from 42.6 ms to 25.1 ms. Participant- and window-level analyses identified significant group differences, particularly during moderate activity. In addition, the proposed ECG quality assessment framework reliably identified high-quality signal segments, achieving near-perfect valid RR ratios (0.99) without manual annotations. Overall, these findings suggest altered, activity-dependent autonomic regulation in cancer survivors and highlight the importance of motion-aware activity segmentation and robust ECG quality control for accurate physiological monitoring in real-world wearable settings.

Index Terms—Activity Intensity, Autonomic Dysfunction, Cancer Survivors, ECG Quality, HRV, Wearable ECG

Manuscript submitted on 1 June 2026.

This work was partially supported by the Florida Department of Health through the Florida Cancer Innovation Fund (Grants MOAAR and 25C38) and by the FIU University Graduate School Dissertation Fellowship. The opinions, findings, and conclusions expressed are those of the authors and do not necessarily reflect the official views of the funding agencies.

S. Farrokhi and C. Poellabauer are with the Knight Foundation School of Computing and Information Sciences at Florida International University, USA, (e-mail: sfarrokh@fiu.edu, cpoellab@fiu.edu)

L. Sequeira and S. Burke are with the School of Social Work at Florida International University, USA, (e-mail: lsequeir@fiu.edu, sburke@fiu.edu)

W. Dargie is with the Faculty of Computer Science, Technische Universität Dresden, 01062 Dresden, Germany (e-mail: waltenegus.dargie@tu-dresden.de)

I. INTRODUCTION

Heart rate (HR) and heart rate variability (HRV) are well-established indicators of autonomic nervous system function, providing critical insights into physiological responses under varying conditions [1]. They are of particular importance for the study of populations whose autonomic regulation may be impaired by specific diseases and medical treatments [2]. Breast Cancer survivors represent one such population.

Wearable electrocardiogram (ECG) technologies enable continuous monitoring of cardiovascular activity in real-world environments [3]. By moving beyond controlled clinical settings, real-world data captures natural variations in activity and stress, offering a richer and more realistic representation of daily physiological behavior. Nevertheless, ECG signals collected in real-world conditions are highly susceptible to motion artifacts and activity-induced variability, which can significantly degrade signal quality. These distortions hinder reliable detection of cardiac events and introduce errors in derived HR and HRV metrics [4], [5]. The problem becomes pronounced when short ECG segments are used for analysis [6]. This issue is further exacerbated during periods of movement or rapidly changing exertion levels. Furthermore, the lack of ground truth annotations in large-scale wearable datasets makes it challenging to assess signal reliability and validate extracted features [7]. Consequently, robust and scalable HRV analysis in real-world environments remains an ongoing challenge.

Knowledge of motion context, signal reliability, and exertion levels is helpful to improve signal selection, analysis, and interpretation. For example, measurements from inertial measurement units (IMUs) can provide reliable estimates of motion intensity [8]. Aligning these measurements with the measurements of wearable electrocardiograms enables the identification of distortions in ECG segments and to reason about underlying causes. However, the state-of-the-art in ECG signal analysis and modeling offers limited insights when it comes to the joint investigation of physiological exertion, ECG signal quality, and heart rate variability [9]. To address these limitations, we propose a unified, quality-aware, and multimodal approach that integrates complementary methods for wearable ECG analysis in free-living environments, with a focus on autonomic regulation in breast cancer survivors. An

overview of the proposed approach is illustrated in Figure 1.

Specifically, we: 1) develop a unified, quality-aware framework that integrates ECG and inertial sensing for activity-aware HR and HRV analysis in real-world environments; 2) propose an unsupervised clustering approach for activity intensity segmentation from accelerometer and gyroscope signals; 3) introduce a clustering-based method for ECG signal quality assessment using signal features; 4) design an annotation-free validation strategy based on physiologically grounded RR interval consistency to identify reliable segments; and 5) evaluate the framework on a real-world dataset of breast cancer survivors and healthy controls, showing that HRV differences were most apparent during moderate physical activity.

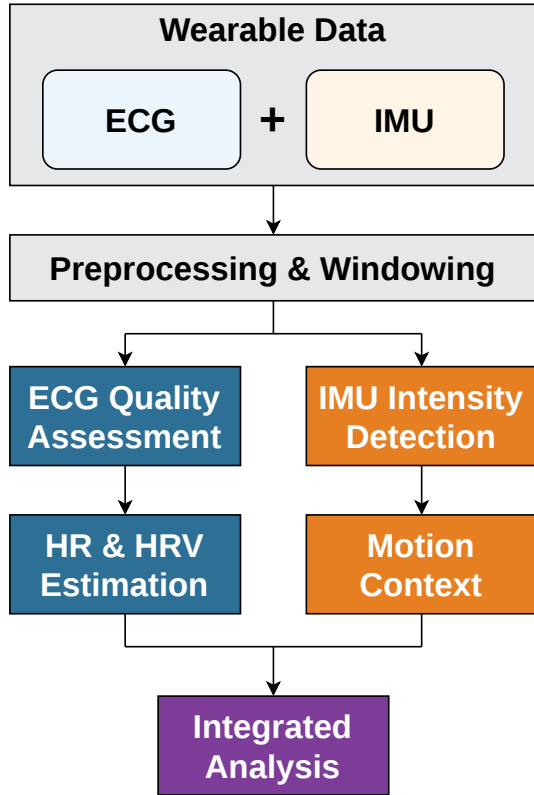


Fig. 1: Overview of the proposed quality-aware multimodal approach for HR and HRV analysis in real-world environments.

The remainder of this paper is organized as follows. Section II reviews related work. Section III presents the proposed methodology, including signal preprocessing, ECG quality assessment, activity intensity segmentation, and statistical analysis. Section IV presents the results and discussion. Finally, Section V concludes the paper.

II. RELATED WORK

A. ECG Signal Degradation and Motion Artifacts in Wearable Settings

Wearable ECG monitoring systems are highly susceptible to signal degradation caused by motion artifacts, environmental interference, and acquisition-related variability. Raimondi et

al. [6] investigated the effect of additive noise on ultra-short-term HRV analysis by injecting controlled noise at different signal-to-noise ratio (SNR) levels, showing that HRV metrics become increasingly unreliable at $\text{SNR} \leq 5$ dB, with stable estimates observed only above 10 dB. Complementing this, Cui et al. [10] analyzed capacitive ECG acquisition by modeling the skin–electrode interface and experimentally evaluating factors such as environmental humidity and clothing thickness. Their results showed that increased skin-electrode impedance attenuates ECG amplitude and reduces waveform reliability, leading to degraded visibility of clinically important ECG features such as QRS complexes.

Similar challenges have been observed in real-world wearable ECG recordings, where motion artifacts caused by body movement, electrode displacement, and muscle activity, along with missing segments and device-dependent sampling inconsistencies, can distort ECG morphology, compromise R-peak detection, and affect downstream HRV analysis [11]. To mitigate these effects, recent studies have explored advanced denoising techniques, including deep learning methods such as generative adversarial networks (GANs) and CycleGAN frameworks. Rajasekhar et al. [12] proposed a CycleGAN-based ECG denoising framework that improved ECG reconstruction quality under noisy conditions while preserving waveform morphology, achieving approximately 2–3 dB higher peak signal-to-noise ratio (PSNR) and lower mean squared error (MSE) compared to conventional GAN approaches. However, these methods are typically evaluated using controlled or simulated noise conditions, and their robustness under highly variable free-living motion artifacts remains less explored.

B. ECG Signal Quality Assessment for Reliable HRV Analysis

To improve the reliability of wearable ECG analysis, previous studies have explored rule-based, learning-based, and unsupervised approaches for ECG signal quality assessment. Clifford et al. [4] introduced signal quality indices (SQIs) based on agreement between multiple QRS detectors and RR interval consistency measures. Although effective in controlled datasets, the approach depends heavily on accurate peak detection and may degrade under severe motion artifacts. More recently, Zhang et al. [13] proposed a framework combining residual and recurrent neural networks for classifying ECG segments into good-, medium-, and low-quality categories, achieving 92.31% accuracy for three-class classification and 98.72% for binary acceptable/unacceptable classification on the CPSC 2020 database containing 87,888 ECG segments, although reduced generalization was observed across heterogeneous populations. Similarly, Yu [14] extracted time- and frequency-domain ECG features and applied an unsupervised K-means clustering framework to separate signal-dominated and noise-dominated segments, improving signal-to-noise ratio up to 21.5 dB while preserving waveform similarity up to 95.8%.

Beyond quantitative noise measures, Holgado-Cuadrado et al. [15] proposed a clinically motivated ECG noise taxonomy

based on waveform interpretability and achieved F1-scores of approximately 0.77 and precision up to 0.80 for distinguishing usable from non-usable ECG segments. Their findings demonstrated that segments with acceptable quantitative noise levels may still lack diagnostically meaningful waveform characteristics. In parallel, previous studies showed that even small inaccuracies in R-peak detection and RR interval estimation can substantially affect downstream HRV metrics, particularly in short-duration wearable recordings [16], [17]. Collectively, these studies demonstrate that while existing methods can assess ECG signal quality effectively, most approaches do not explicitly evaluate the reliability of downstream HR and HRV analysis in free-living wearable environments.

C. HRV, Physical Activity, and Cancer Survivorship

Heart rate variability is strongly influenced by physical activity intensity and autonomic nervous system dynamics. To investigate these effects, Wang *et al.* [18] analyzed HRV responses during progressively increasing exercise intensity using a Bruce treadmill protocol in approximately 23 participants and reported substantial reductions in time-domain and frequency-domain HRV metrics with increasing physiological load. In particular, SDNN decreased from 40.43 ms to 8.70 ms, RMSSD decreased from 39.05 ms to 6.19 ms, and frequency-domain components such as LF and HF decreased by approximately 78%, indicating a shift toward sympathetic nervous system dominance.

In clinical populations, HRV has been widely used to assess autonomic dysfunction and cardiovascular regulation. Several studies have shown that breast cancer survivors exhibit reduced HRV and elevated resting heart rate compared to healthy controls. For example, Caro *et al.* [19] reported reductions in SDNN from 62.35 ± 18.21 ms in controls to 39.10 ± 16.28 ms in survivors and RMSSD from 58.65 ± 22.17 ms to 28.82 ± 23.52 ms. Similarly, Sousa *et al.* [20] observed reduced RMSSD values from 39.15 ± 37.66 ms in controls to 14.89 ± 8.28 ms in breast cancer patients, along with reductions in SDNN from 55.77 ± 40.03 ms to 26.30 ± 10.37 ms. Complementing these findings, Majerova *et al.* [21] analyzed time-domain, frequency-domain, and nonlinear HRV metrics in active cancer patients and survivors, reporting increased sympathetic modulation with 0V% increasing from 16.17% to 26.23% ($p < 0.01$) and reduced signal complexity with sample entropy decreasing from approximately 1.86 to 1.60–1.62 ($p < 0.01$), although RMSSD and SDNN showed limited differences under controlled resting conditions. More broadly, Borsati *et al.* [22] reported that cardiac autonomic dysfunction may affect up to 80% of cancer patients and survivors and associated these alterations with increased sympathetic activation, reduced vagal tone, systemic inflammation, and treatment-related cardiotoxicity.

Despite these findings, most HRV studies in cancer populations have been conducted under controlled resting conditions using short-duration recordings, without accounting for physical activity, posture, or real-world behavioral variability. Similarly, activity-aware HRV studies are typically performed under controlled laboratory conditions where signal quality

is assumed to be reliable and motion-induced artifacts are minimized. As a result, limited work has jointly considered physical activity intensity, wearable ECG signal quality, and HRV estimation in real-world environments. Taken together, these limitations highlight the need for integrated frameworks that jointly consider motion context, ECG signal quality, and reliable HR/HRV estimation under realistic wearable sensing conditions.

III. METHODOLOGY

A. Dataset

The dataset consists of synchronized physiological and motion recordings collected using wearable devices in free-living conditions. Participants were recruited through targeted social media campaigns, and data collection was conducted under an approved Institutional Review Board (IRB) protocol at Florida International University (FIU) over a six-month period in 2025. All participants provided written informed consent prior to participation. The study included female breast cancer survivors aged 18 to 59 who had completed primary cancer treatment within the previous five years and resided in South Florida. The healthy control group consisted of women with no history of cancer. Exclusion criteria for all participants included: (1) conditions impairing motor or cognitive function, (2) inability to independently operate a smartphone application, and (3) inability or unwillingness to use a wearable device. Healthy controls were additionally excluded if they had medical conditions with symptoms similar to dysautonomia, such as endocrine disorders, chronic fatigue syndrome, fibromyalgia, autoimmune disorders, or Ehlers-Danlos syndrome.

Following an initial in-person home visit, participants were trained on device setup, sensor placement, and recording procedures. The sensor was positioned at the center of the chest, with electrode pads moistened prior to placement to ensure optimal signal quality. Participants then conducted recording sessions in free-living environments of their choice, including homes, gyms, and other locations. Each participant completed at least three sessions per week, with a minimum duration of 15 minutes, covering a range of daily activities. Breast cancer survivors were monitored for four weeks, whereas healthy controls were monitored for two weeks to account for the greater expected day-to-day physiological variability in the survivor cohort. In addition to free-living recordings, participants completed a supervised 1-minute seated resting baseline recording during the initial visit to provide a reference measure of resting cardiovascular activity.

Data were acquired using the Movesense MD wearable sensor (Model OP174; Movesense Oy, Espoo, Finland), a Bluetooth Low Energy-enabled Class IIa medical device certified under the European Union Medical Device Regulation (MDR 2017/745) [23]. The sensor records single-lead ECG along with multi-axis IMU and magnetometer signals, enabling joint analysis of cardiac activity and motion. ECG signals were sampled at 125 Hz, while IMU and magnetometer data were sampled at 52 Hz. Participants connected the sensor to their smartphones, and recording sessions were initiated and

TABLE I: Clinical characteristics of the breast cancer survivor cohort ($n = 37$). Unknown indicates unavailable medical records.

Cancer stage	n (%)	Treatment	n (%)	Receptor	n (%)
I	14 (37.8)	Surgery	32 (86.5)	ER+	28 (75.7)
II	13 (35.1)	Chemo	20 (54.1)	PR+	25 (67.6)
III	3 (8.1)	Radiation	10 (27.0)	HER2+	7 (18.9)
IV	1 (2.7)	Hormonal	25 (67.6)	Unknown	2 (5.4)
Unknown	6 (16.2)	Unknown	1 (2.7)		

TABLE II: Dataset size and data retention across processing stages.

Group	N	Total (h)	QC (h)	Final (h)	QC (%)	Final (%)
Survivor	37	272.46	225.71	130.91	82.84	48.05
Control	17	20.09	16.97	11.58	84.50	57.62
Total	54	292.55	242.68	142.49	82.95	48.71

terminated through the Movesense mobile application. In total, data were collected from 54 participants, including 37 breast cancer survivors and 17 healthy controls. The breast cancer survivor group had a mean age of 46.2 ± 7.1 years, mean height of 163.3 ± 5.8 cm, and mean weight of 71.9 ± 15.3 kg, while the healthy control group had a mean age of 26.2 ± 5.2 years, mean height of 163.1 ± 6.9 cm, and mean weight of 63.5 ± 12.7 kg. Clinical characteristics, including disease stage, receptor status, and treatment history, for breast cancer survivors with available medical records are summarized in Table I.

Only segments with concurrent ECG and IMU recordings were retained, resulting in 292.55 hours of data. Due to variations in recording duration and participant compliance, the dataset is imbalanced, with a larger proportion of data from breast cancer survivors. To ensure reliability, the data were further refined through signal quality control and physiological validity constraints based on RR interval availability and HRV plausibility. After filtering, 142.49 hours of data remained for analysis, comprising 130.91 hours from the breast cancer survivor group and 11.58 hours from the control group. Table II summarizes the dataset size and data retention across processing stages.

B. Signal Preprocessing

The ECG signal was preprocessed using zero-phase Butterworth bandpass filtering, implemented via forward-backward filtering to avoid phase distortion [24]. The filter range was adapted to the analysis task. For signal characterization, clustering, and feature extraction, a 1–40 Hz band was applied to retain physiologically relevant ECG components while suppressing baseline drift and high-frequency noise, consistent with standard ECG preprocessing practices [16]. In contrast, for R-peak detection, which forms the basis for subsequent HRV computation, a wider band of 1–60 Hz was used to preserve higher-frequency components of the QRS complex and improve peak localization accuracy under motion artifacts.

For IMU signals, accelerometer measurements were scaled when necessary to ensure consistent physical units. The gravitational component was removed using high-pass filtering with a cutoff frequency of 0.25 Hz, attenuating low-frequency

TABLE III: ECG features used for signal quality assessment.

Category	Features
Amplitude	BP STD (1–40 Hz), RMS
Spectral	Low-band ratio, HF ratio
Morphology	Derivative RMS

Abbreviations: BP = bandpass; STD = standard deviation; RMS = root mean square; HF = high frequency.

components associated with device orientation and isolating dynamic acceleration due to motion, as commonly performed in wearable sensing applications [25]. For clustering, extracted features were standardized using z-score normalization ($x_{\text{norm}} = (x - \mu)/\sigma$) to zero mean and unit variance.

Following preprocessing, both ECG and IMU signals were segmented into overlapping windows of 10 s with a 5 s step, enabling short-term analysis of physiological and motion dynamics while maintaining sufficient temporal resolution for capturing activity transitions. Additional task-specific preprocessing steps are described within the corresponding subsections.

C. ECG Signal Quality Validation Without Annotations

In real-world environments, ECG signals are frequently affected by motion artifacts and signal distortions, making reliable identification of physiologically meaningful segments challenging in the absence of manual annotations [3]. To address this, we adopt a data-driven approach in which ECG signal quality is first inferred through unsupervised clustering based on extracted ECG signal features. These features, summarized in Table III, capture amplitude, spectral, and morphological characteristics commonly used in ECG signal analysis [26]. The resulting clusters are subsequently evaluated using physiologically grounded metrics derived from R-peak detection, which are not used in forming the clusters.

Specifically, ECG signals were segmented into overlapping 10 s windows (5 s step) and grouped using unsupervised clustering based on signal characteristics. The resulting clusters are hypothesized to represent latent ECG signal quality levels. To evaluate clustering performance, standard validation metrics were computed, including the silhouette score [27] and Davies–Bouldin index [28]. The clustering achieved a silhouette score of 0.56 and a Davies–Bouldin index of 0.60, indicating well-defined and compact cluster structure, thereby supporting the validity of the derived ECG quality groups. However, because clustering is performed without supervision, an additional validation step is required to ensure that these groupings correspond to meaningful differences in physiological signal reliability. To this end, R-peaks were detected within each ECG window using the method described in Section III-E, and the detected peaks were used to derive a set of annotation-free metrics that assess the physiological plausibility and consistency of the resulting RR intervals.

The reliability of RR intervals was first evaluated based on their physiological range. The *valid RR ratio* was defined as the proportion of intervals within 300–2000ms, consistent with established HRV preprocessing practices [16], [29], and serving as a primary indicator of reliable peak detection. Temporal

consistency was further assessed using the *RR instability ratio*, defined as the proportion of consecutive RR interval ratios falling outside predefined bounds (0.60–1.67). This metric evaluates how much the heart rate is allowed to change from one beat to the next. A ratio of 1 indicates no change, while values below 0.60 correspond to a sudden shortening of the RR interval by more than 40%, and values above 1.67 correspond to a sudden lengthening by more than 67%. Such large beat-to-beat changes are unlikely under normal physiological conditions, where heart rate typically varies gradually. Therefore, ratios outside this range indicate abrupt and non-physiological fluctuations, which are commonly associated with missed or spurious peak detections [17]. To further characterize signal stability, the *coefficient of variation of RR intervals (RR CV)* was computed for each window, where higher values indicate increased variability and reduced signal reliability. In addition to temporal and statistical measures of RR variability, morphological consistency was subsequently assessed using *beat-template correlation*, where individual beats were aligned and compared to a local template, and the average correlation was used as a measure of waveform consistency [7].

These metrics, summarized in Table IV, were used specifically for physiological validation of the ECG quality clusters rather than as outcome measures in the subsequent HRV analysis. They were aggregated within each cluster to evaluate whether clustering effectively separates ECG segments according to physiological signal reliability. Clusters exhibiting high valid RR ratios, low variability, and high morphological consistency were interpreted as *high-quality* ECG segments suitable for HRV analysis.

It is important to note that these labels do not directly correspond to noise levels, but rather reflect the reliability and physiological plausibility of the detected cardiac activity within each cluster. As illustrated in Figure 2, the clustered ECG windows exhibit clear separation in the feature space, supporting the assumption that unsupervised clustering captures distinct signal quality regimes.

As an additional validation step, each 10 s ECG window was independently assigned a reliability label (high, moderate, or low) using the proposed quality metrics. The default thresholds were a valid RR ratio ≥ 0.90 , RR instability ratio ≤ 0.15 , RR CV ≤ 0.10 , and beat-template correlation ≥ 0.70 . These thresholds were empirically selected to balance physiological plausibility and signal consistency with the variability expected in free-living ECG recordings. Windows satisfying the majority of these criteria were labeled as reliable, those satisfying a subset as moderately reliable, and those failing most criteria as poorly reliable. To assess the robustness of these empirically selected thresholds, the analysis was repeated using permissive and strict configurations. The permissive thresholds were 0.85, 0.20, 0.15, and 0.65, respectively, whereas the strict thresholds were 0.95, 0.10, 0.05, and 0.75.

The distribution of the independently assigned reliability labels across clusters is reported in Table VI. Under the default configuration, 96.1% of the windows in the high-quality cluster were independently classified as reliable, compared with 20.4% and 19.8% in the moderate- and low-quality clusters, respectively. Under the permissive and strict threshold

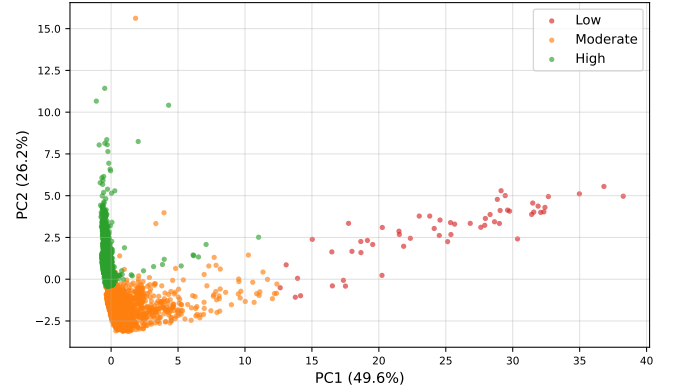


Fig. 2: Projection of ECG windows onto the first two principal components (PC1 and PC2), colored by cluster assignment. The high-quality cluster forms a compact, well-defined region, whereas the lower-quality clusters are more dispersed, reflecting greater signal variability.

TABLE IV: Annotation-free ECG quality metrics used for cluster validation.

Metric	Description
Valid RR ratio	Fraction of RR intervals within 300–2000ms
RR instability ratio	Fraction of abrupt changes between consecutive RR intervals
RR CV	Coefficient of variation of RR intervals within each window
Beat-template correlation	Average correlation between individual beats and a local template

configurations, the proportion of reliable windows in the high-quality cluster changed only modestly (97.2% and 95.3% for cancer survivors; 92.0% and 86.6% for healthy controls), while it remained the cluster with the highest proportion of reliable windows across all configurations. These findings indicate that the identification of the high-quality ECG cluster was robust to reasonable variations in the selected thresholds. In contrast, the moderate- and low-quality clusters exhibited more heterogeneous reliability-label distributions, reflecting gradual transitions and increased variability in signal reliability.

The results support the conclusion that clustering-based grouping aligns with physiological signal quality. As shown in Table V, the high-quality cluster achieves near-perfect RR validity and strong waveform consistency, whereas lower-quality clusters exhibit substantially degraded metrics. Because cluster quality was determined using the combined performance across all validation metrics, rather than any single metric individually, the assigned quality labels reflect the overall physiological consistency of each cluster. This indicates that the proposed approach effectively identifies reliable ECG segments without requiring manual annotations, enabling robust HRV analysis in real-world settings. Figure 3 provides representative examples of ECG segments across different quality levels alongside the corresponding activity intensity measurements.

TABLE V: Annotation-free ECG quality metrics across clusters (mean values).

Cluster	Quality	Valid RR	RR CV	Beat Corr	Windows (%)
High		0.99	0.02	0.94	63.8
Moderate		0.40	0.25	0.82	35.6
Low		0.88	0.33	0.61	0.6

TABLE VI: Distribution of ECG signal reliability labels across clusters. Values are reported as percentages.

Cluster	Quality	High Reliability	Moderate Reliability	Low Reliability
High		96.1	3.1	0.8
Moderate		20.4	43.1	36.5
Low		19.8	39.1	41.1

D. Clustering-Based Activity Intensity Segmentation

To characterize activity intensity from wearable sensor data, a clustering-based segmentation approach was applied to accelerometer (ACC) and gyroscope (GYRO) signals. Data were segmented into overlapping windows of 10 s with a 5 s step, enabling short-term characterization of motion patterns in free-living conditions.

For each modality, window-level features were extracted to capture signal properties relevant to activity intensity (see Table VII). These features characterize signal magnitude, variability, and motion dynamics, including vector magnitude and Euclidean Norm Minus One (ENMO), which are widely used indicators of physical activity intensity [8]. All features were first standardized using z-score normalization to ensure comparability across dimensions. Global K-means clustering with $K = 3$ was then applied to ACC and GYRO features to partition windows into three groups [30]. To improve clustering stability and reproducibility, 20 random initializations were performed using a fixed random seed (42), and the solution with the lowest within-cluster sum of squares was retained. The resulting clusters were ordered based on signal-energy measures and mapped to clustering-derived *light*, *moderate*, and *vigorous* activity intensity categories. These clustering-derived labels represent relative motion intensity and should not be interpreted as standardized physiological intensity categories. Figure 4 shows a clear monotonic increase in vector magnitude RMS across the derived intensity groups.

Figure 5 further illustrates the modality-specific structure of the derived groups, where ACC and GYRO feature spaces exhibit clear separation across activity levels. Clustering performance was evaluated using the Silhouette score [27] and Davies–Bouldin index [28]. As shown in Table VIII, both IMU modalities achieve strong separation, supporting the effectiveness of the proposed segmentation approach. Finally, modality-specific segmentations were combined by assigning each window the higher intensity level derived from ACC and GYRO signals, ensuring that periods of increased movement detected by either sensor were retained in the final labeling and reducing the risk of underestimating activity intensity.

Although clustering was performed using $K = 3$ to capture the full spectrum of activity intensities, the vigorous activity group was excluded from downstream statistical analyses because vigorous activity is substantially more susceptible

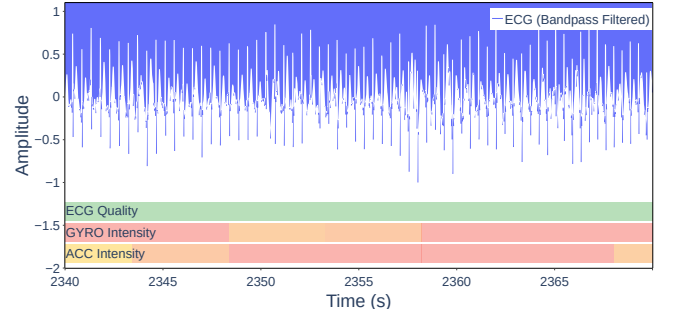
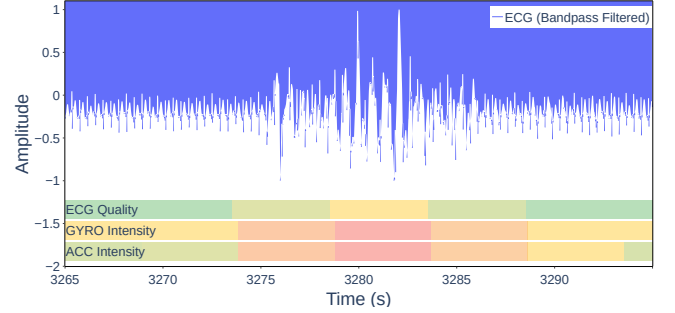
**(a)** High-intensity segment with high-quality ECG**(b)** High-intensity segment with moderate-quality ECG

Fig. 3: Representative ECG segments recorded during high-intensity activity. Despite similar motion intensity, the segments differ in ECG quality. The top panel shows a high-quality segment with stable morphology, whereas the bottom panel shows a moderate-quality segment with increased variability. Colored bands indicate ECG quality (green: high, yellow: moderate) and IMU-derived activity intensity (green: low, yellow: moderate, red: high).

TABLE VII: ACC and GYRO feature groups used for activity intensity classification.

Category	Features
Magnitude	VM mean, RMS
Energy	ENMO mean, ENMO RMS, jerk RMS
Variability	MAD, axis-wise STD
Frequency	Step-band ratio, tilt variance

Abbreviations: VM (vector magnitude); ENMO (Euclidean norm minus one); MAD (mean absolute deviation); BP (bandpass); HF (high frequency); STD (standard deviation).

to motion-related artifacts, resulting in reduced ECG signal reliability and less robust HR and HRV estimation. In addition, the vigorous activity group represented only approximately 3.3% of the ECG windows after quality filtering, limiting the statistical power for meaningful comparisons. Consequently, subsequent HR and HRV analyses focused on light and moderate activity intensities, where both ECG signal quality and sample representation were more suitable for reliable statistical analysis.

E. Peak Detection

Following preprocessing and signal quality validation, physiological features were extracted from ECG signals using a sliding window framework. R-peaks were detected within

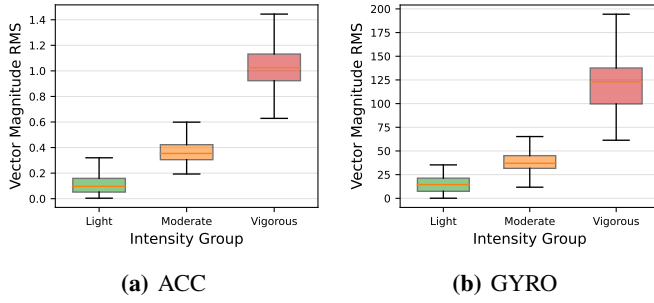


Fig. 4: Distribution of vector magnitude RMS (VM RMS) across clustering-derived activity intensity groups, showing a monotonic increase from light to vigorous activity.

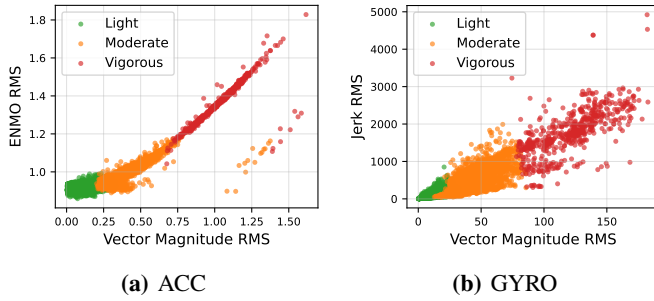


Fig. 5: Feature-space representation of clustering-derived activity intensity groups from accelerometer (ACC) and gyroscope (GYRO) signals. Colors indicate cluster assignments.

10 s ECG windows using an extended version of our previously proposed framework [31]. The method enhances the QRS complex through bandpass filtering and wavelet-based decomposition, enabling reliable identification of candidate peaks even in the presence of noise. Peak candidates are then identified using adaptive amplitude thresholding, where detection thresholds are dynamically adjusted based on segment-level signal characteristics. To ensure robustness under motion artifacts, the initial R-peak detections were evaluated based on temporal consistency, where windows exhibiting implausible or highly irregular RR intervals were identified as unreliable. In such cases, an adaptive fallback mechanism [32] was triggered to re-estimate R-peaks using relaxed detection thresholds and temporal constraints, enabling recovery from missed or spurious detections. Following re-estimation, peak locations were refined by aligning each detected peak to the nearest local maximum within a small neighborhood to improve temporal precision.

RR intervals were computed from the resulting R-peak sequence and subsequently aggregated over analysis windows of 30 s, 1 min, and 2 min. Although higher ECG sampling rates may improve R-peak localization accuracy, the 125 Hz sampling rate used in this study has been shown to provide reliable estimation of time-domain HRV metrics when accurate R-peak detection is employed [33]. No interpolation or upsampling was applied to the ECG signal. The resulting RR intervals then underwent a final temporal plausibility check, ensuring that they remained within physiologically valid ranges and did not exhibit abrupt, non-physiological beat-to-beat variations.

TABLE VIII: Activity intensity detection using motion sensors with ECG-based physiological validation.

Sensor	Silhouette Score	Davies–Bouldin Score
ACC	0.43	0.80
GYRO	0.31	1.02

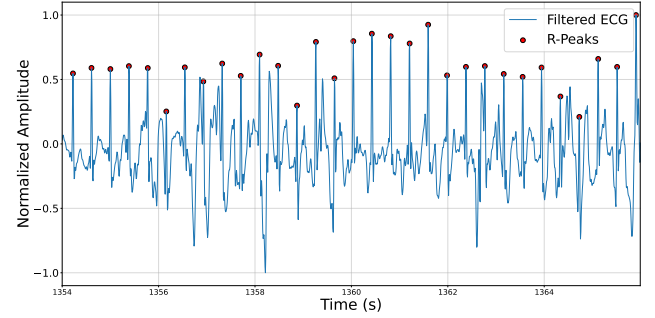


Fig. 6: ECG signal with detected R peaks

Segments failing prior signal quality criteria or violating these constraints were excluded from further analysis. Figure 6 illustrates an example of the detected R-peaks on a representative ECG segment.

F. Statistical Analysis

To investigate physiological differences between cancer survivors and healthy controls, heart rate and heart rate variability metrics were analyzed across light and moderate activity intensities. HRV was quantified using the root mean square of successive RR interval differences (RMSSD) and the standard deviation of quality-controlled RR intervals (SDRR). The RR intervals were obtained following rigorous ECG signal quality assessment, correction of suspected missed and spurious R-peak detections, and physiological plausibility screening. Because explicit ectopic beat detection and correction were not performed, the latter metric is reported as SDRR rather than SDNN, consistent with the standard definition of SDNN based on normal-to-normal intervals [34]. Nevertheless, because SDRR is computed using the same standard deviation formulation from rigorously quality-controlled RR intervals, qualitative comparisons with prior studies reporting SDNN are reasonable while maintaining technically accurate terminology. These metrics were computed using sliding windows of 30 s, 1 min, and 2 min to capture cardiovascular responses across multiple temporal scales.

To ensure statistically robust and valid comparisons, analyses were conducted at both the participant and window levels. At the participant level, HR and HRV measurements were aggregated within each individual to preserve independence between observations and avoid bias from repeated measures. Group differences were then evaluated using Welch's t-test, which is appropriate under unequal variances and sample sizes [35]. As a participant-level sensitivity analysis, RMSSD and SDRR were additionally normalized by the corresponding mean RR interval for each 1-min analysis window (RMSSD/mean RR and SDRR/mean RR, respectively), and participant-level median values of these normalized metrics

were compared between groups using the same statistical procedures as the primary analysis. To further account for repeated measurements while adjusting for potential confounding factors, linear mixed-effects models were fitted using participant-level median 1-min HR, RMSSD, and SDRR measurements obtained during light and moderate activity. Group (breast cancer survivor versus healthy control), activity intensity, and age were included as fixed effects, with participant modeled as a random intercept [36]. HR was analyzed on its original scale, whereas RMSSD and SDRR were natural log-transformed.

To further capture temporal dynamics, a complementary window-level analysis was performed in which each valid window was treated as an observation. Given the potential for non-normal distributions, heteroscedasticity, and imbalanced sample sizes at this level, both parametric and nonparametric tests were applied. Specifically, Welch's t-test was used to assess differences in group means, while the Mann-Whitney U test was used to evaluate distributional differences without assuming normality [37]. To account for multiple comparisons across the HRV analyses, Welch's p -values were adjusted using the Benjamini-Hochberg false discovery rate (FDR) procedure [38]. The correction was applied across all participant- and window-level comparisons for RMSSD and SDRR (24 tests in total). Raw and FDR-adjusted p -values are reported for the participant-level analyses in Table X. Statistical significance was defined as $p < 0.05$. In addition to hypothesis testing, effect sizes were quantified using Cohen's d to assess the magnitude of group differences independent of sample size [39]. This is particularly important in window-level analyses, where large numbers of observations may lead to statistically significant results with limited practical relevance. Accordingly, both statistical significance and effect size are considered when interpreting the results.

Overall, this multi-level statistical framework enables a comprehensive and robust evaluation of differences in HR and HRV between cancer survivors and healthy controls across varying activity intensities and temporal resolutions, while accounting for common challenges in wearable physiological data, including non-normality, unequal variances, and sample size imbalance.

IV. RESULTS AND DISCUSSION

This section examines differences in cardiac responses to physiological exertion between cancer survivors and healthy controls. We first analyze heart rate and heart rate variability across activity intensities and window lengths, followed by participant- and window-level statistical comparisons.

A. Heart Rate and HRV Across Activity Intensities

Table IX presents the mean HR and HRV values for both groups across activity intensities. Baseline measurements were available only for the 1 min analysis windows and are included for reference. As expected, HR increased with activity intensity in both groups, reflecting normal cardiovascular responses to increased physical demand. However, cancer survivors consistently exhibited higher HR values during light

activity compared to controls. For example, in the 30 s windows, cancer survivors showed an average HR of 103.44 bpm compared to 95.70 bpm for controls. Figure 7 illustrates the corresponding HR distributions for the 30 s windows across activity intensities for both groups. In contrast, HRV metrics were consistently lower in cancer survivors across most activity intensities and window lengths. These differences were particularly pronounced during moderate activity. For instance, in the 1 min windows during moderate activity, RMSSD averaged 22.09 ms for cancer survivors compared to 39.68 ms for controls, while SDRR was 25.14 ms for cancer survivors and 42.62 ms for controls. Similar patterns were observed across other window lengths.

B. Participant-Level Statistical Comparisons

To assess participant-level differences between cancer survivors and healthy controls, Welch's t-tests were conducted, as summarized in Table X. The most consistent group differences were observed during moderate activity.

In the age-unadjusted participant-level analysis, moderate activity yielded statistically significant differences in RMSSD across all window lengths (30 s: $p = 0.04$, $d = -0.63$; 1 min: $p = 0.02$, $d = -0.75$; 2 min: $p = 0.019$, $d = -0.67$), with moderate effect sizes. A similar pattern was observed for SDRR, with significant differences during moderate activity (30 s: $p = 0.02$, $d = -0.73$; 1 min: $p = 0.004$, $d = -1.03$; 2 min: $p = 0.003$, $d = -0.98$), where effect sizes were generally larger for SDRR than RMSSD, particularly during light and moderate activity. In contrast, differences during light activity were not statistically significant for either metric across window lengths, with small effect sizes ($|d| \leq 0.45$), indicating minimal group differences under low physical demand. Following Benjamini-Hochberg FDR correction for multiple comparisons (24 HRV comparisons), the overall findings remained largely unchanged, with significant differences retained for SDRR during moderate activity across all window lengths and for RMSSD during moderate activity for the 1- and 2-min windows. As a participant-level sensitivity analysis, the participant-level comparisons were repeated using HR-normalized HRV metrics. The overall pattern of group differences remained after normalization. During moderate activity, cancer survivors exhibited lower RMSSD/mean RR (0.032 ± 0.031 vs. 0.066 ± 0.053 , $p = 0.030$, $d = -0.84$) and SDRR/mean RR (0.036 ± 0.020 vs. 0.072 ± 0.033 , $p < 0.001$, $d = -1.44$) than healthy controls. During light activity, SDRR/mean RR also remained significantly lower in cancer survivors (0.044 ± 0.026 vs. 0.060 ± 0.025 , $p = 0.048$, $d = -0.65$), whereas RMSSD/mean RR did not differ significantly between groups (0.038 ± 0.036 vs. 0.049 ± 0.028 , $p = 0.275$, $d = -0.32$). These findings suggest that the reduced HRV observed in cancer survivors is not solely attributable to differences in mean heart rate.

To further assess the influence of repeated measurements and age, age-adjusted linear mixed-effects models were fitted using the 1-min measurements (Table XI). Although adjustment for age and repeated measurements attenuated the estimated group effects, the direction of the associations remained consistent with the primary participant-level analysis,

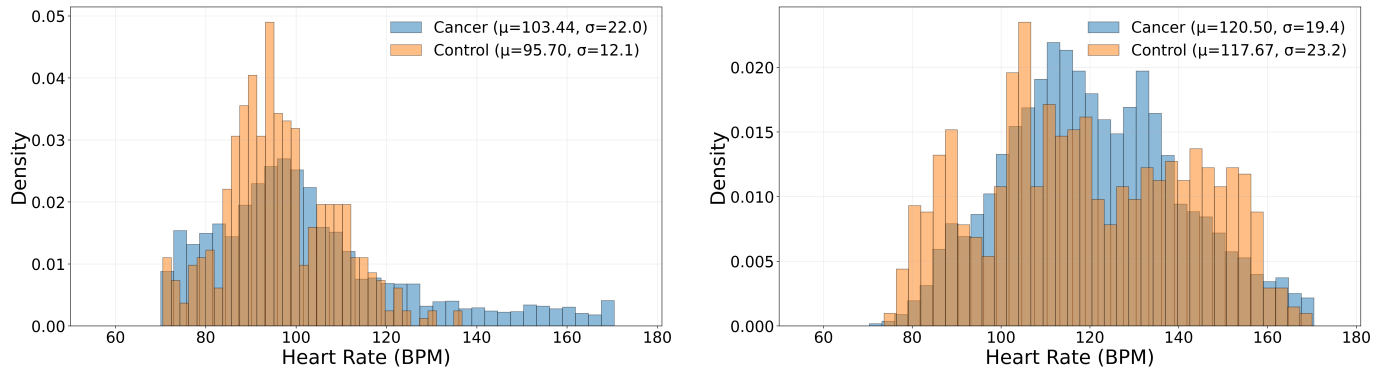


Fig. 7: Heart rate distributions for cancer survivors and healthy controls across light and moderate activity intensities. Cancer survivors exhibit elevated heart rates and broader variability across activity levels. Window-level statistical comparisons showed significant differences during light activity ($p \approx 1 \times 10^{-22}$) and moderate activity ($p \approx 0.0017$).

TABLE IX: Mean HR (bpm), RMSSD (ms), and SDRR (ms) for cancer survivors and healthy controls across activity intensities and window lengths. Baseline values are available only for the 1-min window.

Window	Group	HR Mean			RMSSD Mean			SDRR Mean		
		Baseline	Light	Moderate	Baseline	Light	Moderate	Baseline	Light	Moderate
30 s	Cancer	–	103.44	120.50	–	27.90	26.18	–	30.08	25.37
	Control	–	95.70	117.67	–	35.67	56.56	–	44.32	50.09
1 min	Cancer	79.90	100.26	119.77	44.50	25.32	22.09	52.71	30.66	25.14
	Control	83.03	93.82	116.14	43.47	31.58	39.68	58.50	44.01	42.62
2 min	Cancer	–	98.00	118.86	–	28.85	25.57	–	35.69	30.67
	Control	–	91.89	114.20	–	31.69	45.34	–	47.34	48.89

with cancer survivors exhibiting higher HR and lower HRV than healthy controls. None of the age-adjusted group effects reached statistical significance (HR: $p = 0.17$; $\ln(\text{RMSSD})$: $p = 0.075$; $\ln(\text{SDRR})$: $p = 0.11$). These findings suggest that age may explain a substantial portion of the observed between-group differences. However, because age and group membership were strongly associated and the cohorts had limited age overlap, the independent effect of cancer survivorship could not be reliably estimated in the present sample. Additionally, free-living HRV measurements may be influenced by uncontrolled factors such as respiratory rate, posture, speech, and emotional stress, which were not standardized in this study.

C. Window-Level Statistical Comparisons

Window-level analysis further characterizes group differences across activity intensities. Across window sizes, HRV metrics were generally lower in cancer survivors compared to healthy controls, with the magnitude of differences varying by activity intensity.

Both Welch's t -test and the Mann–Whitney U test indicated statistically significant differences between groups for nearly all comparisons ($p < 0.05$). All window-level comparisons that were statistically significant before correction remained significant after Benjamini–Hochberg FDR adjustment. However, due to the large number of window-level observations and repeated measurements within participants, statistical significance is expected and may be inflated. Therefore, interpretation focuses primarily on effect sizes, which provide a more meaningful measure of group differences.

TABLE X: Participant-level comparison of HRV metrics between cancer survivors and healthy controls using Welch's t -test. Here, t denotes the Welch's t statistic, p the p -value, FDR the Benjamini–Hochberg false discovery rate adjusted p -value (24-test family), d Cohen's effect size, and CI the bootstrap 95% confidence interval for d . Statistically significant results ($p < 0.05$ or FDR-adjusted $q < 0.05$) are shown in bold.

Window	Metric	Intensity	t	p	FDR	d	95% CI
30 s	RMSSD	Light	-0.41	0.68	0.68	-0.11	(-0.97, 0.25)
		Moderate	-2.11	0.04	0.056	-0.63	(-1.60, -0.02)
	SDRR	Light	-0.92	0.36	0.39	-0.33	(-1.25, 0.03)
		Moderate	-2.43	0.02	0.03	-0.73	(-2.04, -0.43)
1 min	RMSSD	Light	-1.04	0.31	0.35	-0.30	(-1.08, 0.23)
		Moderate	-2.43	0.02	0.03	-0.75	(-1.74, -0.15)
	SDRR	Light	-1.09	0.28	0.35	-0.32	(-1.14, 0.21)
		Moderate	-3.15	0.004	0.007	-1.03	(-1.90, -0.40)
2 min	RMSSD	Light	-0.88	0.40	0.42	-0.36	(-1.50, 0.32)
		Moderate	-2.41	0.019	0.03	-0.67	(-1.51, -0.12)
	SDRR	Light	-1.06	0.31	0.35	-0.45	(-1.45, 0.29)
		Moderate	-3.16	0.003	0.006	-0.98	(-1.73, -0.39)

As illustrated in Figure 8, effect size analysis highlights consistent activity-dependent differences between groups. During light activity, effect sizes are negligible to small (e.g., RMSSD $d = -0.20$ for 30 s windows, $d = -0.13$ for 1 min windows, and $d = -0.04$ for 2 min windows), indicating weak separation between groups under low physical demand. In contrast, moderate activity produces consistently moderate-to-large effect sizes (e.g., RMSSD $d = -0.72$, $d = -0.67$, and $d = -0.63$ for 30 s, 1 min, and 2 min windows, respectively). SDRR follows the same trend, with small differences during light activity (e.g., $d = -0.47$, $d = -0.37$, and $d = -0.31$)

TABLE XI: Linear mixed-effects model estimates for the 1-min measurements. Fixed effects included group and activity intensity, with age added in the adjusted model. Participant was modeled as a random intercept. Healthy controls and light activity were the reference categories.

Model	Outcome	β (95% CI)	p
Unadjusted	HR	-0.16 (-9.91, 9.58)	0.974
	ln(RMSSD)	-0.53 (-1.02, -0.03)	0.037
	ln(SDRR)	-0.73 (-1.12, -0.35)	< 0.001
Age-adjusted	HR	9.36 (-4.02, 22.74)	0.170
	ln(RMSSD)	-0.586 (-1.24, 0.07)	0.075
	ln(SDRR)	-0.395 (-0.91, 0.12)	0.110

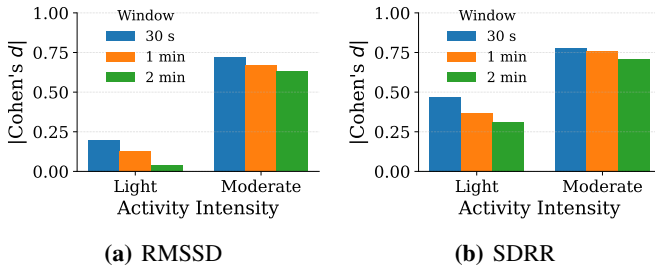


Fig. 8: Magnitude of effect sizes ($|Cohen's d|$) across activity intensities and window lengths. Effect sizes were larger during moderate than light activity, indicating greater separation between breast cancer survivors and healthy controls.

and consistently larger effects during moderate activity (e.g., $d = -0.78$, $d = -0.76$, and $d = -0.71$). Consistent with the participant-level analysis, SDRR generally exhibits larger effect sizes than RMSSD during light and moderate activity. Detailed window-level statistics, including the number of observations per group, are reported in Table XII for completeness.

D. Discussion and Comparison with Prior Work

In the age-unadjusted analyses, the strongest and most consistent HRV differences were observed during moderate activity. As shown in Table IX, cancer survivors exhibit lower HRV compared to controls across most activity conditions and window lengths. For example, in 1-minute windows during moderate activity, RMSSD decreases from 39.68 ms in controls to 22.09 ms in cancer survivors, while SDRR decreases from 42.62 ms to 25.14 ms. These reductions persist across 30-second and 2-minute windows, indicating that the observed group differences are robust to temporal resolution.

These findings are consistent with prior clinical studies reporting altered autonomic regulation in cancer populations. Sousa et al. [20] conducted a prospective case-control study with pre- and post-chemotherapy assessments under controlled resting conditions, evaluating HRV using standard time- and frequency-domain metrics. Their results demonstrated significant reductions in HRV following chemotherapy, with RMSSD decreasing from 44.43 ± 25.27 ms to 14.89 ± 8.28 ms and SDNN from 72.14 ± 61.37 ms to 26.30 ± 10.37 ms in the cancer group, while the control group showed opposite trends,

TABLE XII: Window-level HRV effect sizes (Cohen's d) between cancer survivors and healthy controls across window lengths and activity intensities.

Window	Metric	Intensity	Cancer n	Control n	d	ES
30 s	RMSSD	Light	4147	458	-0.20	S
		Moderate	11063	869	-0.72	M-L
	SDRR	Light	4147	458	-0.47	S-M
		Moderate	11063	869	-0.78	L
1 min	RMSSD	Light	1479	229	-0.13	S
		Moderate	6219	660	-0.67	M
	SDRR	Light	1479	229	-0.37	S
		Moderate	6219	660	-0.76	L
2 min	RMSSD	Light	492	167	-0.04	N
		Moderate	3430	550	-0.63	M
	SDRR	Light	492	167	-0.31	S
		Moderate	3430	550	-0.71	M-L

ES: effect size (N: negligible, S: small, M: moderate, L: large).

indicating impaired parasympathetic modulation and overall autonomic imbalance.

Similarly, Majerova et al. [21] evaluated HRV using 5-minute resting recordings in a controlled seated environment, analyzing time-domain, frequency-domain, and nonlinear metrics. While their time-domain analysis did not reveal significant differences in RMSSD and SDNN between groups, nonlinear measures identified increased sympathetic modulation in cancer survivors, reflected by higher $0V\%$ values ($16.17 \pm 6.44\%$ vs. $26.23 \pm 9.7\%$) and reduced signal complexity indicated by lower sample entropy (1.86 ± 0.21 vs. $1.60 - 1.62$). The absence of significant differences in conventional time-domain metrics suggests that resting HRV analysis may not fully capture HRV differences in this population. This is consistent with the present findings, where baseline HRV values in the 1-minute windows were relatively comparable between groups (RMSSD: 44.50 ms vs. 43.47 ms; SDRR: 52.71 ms vs. 58.50 ms), while substantially larger differences emerged during moderate activity.

The present results reconcile these findings by showing that differences in RMSSD and SDRR were more apparent during moderate activity than at baseline. **This pattern suggests that activity context may influence the magnitude of observed HRV differences.** Although adjustment for age weakened the observed group differences, the limited age overlap between the survivor and control cohorts makes it difficult to fully separate the effects of age and cancer survivorship. Previous studies in healthy populations show variability in the magnitude of age-related HRV differences. Zhang et al. reported an approximately 18% reduction in RMSSD between participants in their 20s and 40s (43.8 vs. 36.0 ms), while SDNN decreased from approximately 55 to 45 ms across the same age groups; the overall effect of age was significant for SDNN but not for RMSSD ($p = 0.229$) [40]. Voss et al. reported reductions of approximately 39% in RMSSD (42.9 vs. 26.3 ms) and 24% in SDNN (48.7 vs. 36.9 ms) between healthy women aged 25–34 and 45–54 [41]. In comparison, the present study observed reductions of approximately 44% in RMSSD (39.68 vs. 22.09 ms) and 41% in SDRR (42.62 vs. 25.14 ms) during moderate activity. Importantly, these previ-

ous studies evaluated HRV under resting conditions, whereas the present study examined 1-min windows during free-living physical activity. Thus, although age likely contributed to the observed HRV differences, the limited age overlap prevents the relative contributions of age and cancer survivorship from being reliably separated.

Beyond baseline differences, a key observation of this study is the distinct manner in which HRV responds to increasing activity intensity across groups. In healthy controls, short-term variability increases from light to moderate activity, as reflected by an increase in RMSSD from 31.58 ms to 39.68 ms, while SDRR remains relatively stable (44.01 ms to 42.62 ms). In contrast, cancer survivors showed decreases in both RMSSD and SDRR from light to moderate activity (RMSSD: 25.32 ms to 22.09 ms; SDRR: 30.66 ms to 25.14 ms). This behavior contrasts with findings from controlled exercise studies, where HRV typically decreases following high-intensity exertion due to increased sympathetic activation. For example, Wang *et al.* [18] evaluated HRV using a structured treadmill-based Bruce protocol, with measurements obtained from 5-minute quiet-sitting recordings before exercise and during recovery after exhaustive exertion. Their results showed marked reductions in HRV, with RMSSD decreasing from 39.05 ms to 6.19 ms and SDNN from 40.43 ms to 8.70 ms ($p < 0.01$), indicating a shift toward sympathetic dominance. Unlike such controlled protocols with fixed pacing and continuous exertion, where HRV was quantified before and after exercise, the present study analyzed HRV continuously during ongoing free-living daily activities, with activity intensity derived from IMU-based measurements rather than prescribed exercise stages. Frequent changes in pacing and brief recovery periods may contribute to greater short-term variability in HRV measurements during moderate activity, despite restricting the analysis to ECG segments classified as high quality. Consequently, the activity intensity categories used in this study are not directly comparable to laboratory-defined exercise intensities.

Despite these methodological differences, cancer survivors exhibited reductions in both RMSSD and SDRR with increasing activity intensity, suggesting different activity-dependent HRV responses during real-world daily activities. An additional observation is the influence of window length on HRV estimation. Shorter windows (30 s) provide greater temporal resolution but contain fewer RR intervals and may therefore produce more variable HRV estimates, particularly for SDRR. For example, RMSSD in the control group during moderate activity was 56.56 ms in 30-second windows, compared with 39.68 ms and 45.34 ms in 1- and 2-minute windows, respectively. While longer windows generally improve estimate stability, they may attenuate short-term dynamics or span changes in activity or physiological state, highlighting a trade-off between temporal resolution and statistical reliability. In this study, 1-minute windows provide a practical balance for capturing dynamic HRV behavior in free-living conditions.

From a physiological perspective, the lower HRV observed in cancer survivors is consistent with patterns associated with increased sympathetic activation and reduced parasympathetic modulation. Potential contributing mechanisms in-

clude chronic inflammation, neuroendocrine dysregulation, treatment-related cardiotoxicity, and surgical stress responses. However, these mechanisms were not directly measured in the present study, and additional studies are needed to determine their contribution to the observed differences. From an application perspective, these findings have important implications for real-world physiological monitoring. The results suggest that reliable assessment of autonomic function in wearable settings requires not only robust signal processing, but also explicit consideration of activity context and signal quality. In particular, the largest group differences were observed during moderate activity, suggesting that this activity level may provide a useful context for assessing autonomic function in wearable monitoring systems. This has potential applications in remote patient monitoring, early detection of cardiovascular dysregulation, and personalized health assessment in clinical populations.

Overall, these findings suggest activity-dependent HRV differences between breast cancer survivors and healthy controls, highlighting the importance of activity-aware HRV analysis for physiological monitoring in real-world wearable environments.

V. CONCLUSION

This study presents a unified, quality-aware multimodal framework for analyzing heart rate and heart rate variability in free-living environments using wearable ECG and inertial sensor data. By integrating activity intensity segmentation, ECG signal quality assessment, and physiological feature extraction, the proposed approach enables reliable analysis of cardiovascular dynamics under real-world conditions. Breast cancer survivors generally exhibited elevated HR and reduced HRV compared to healthy controls, with the largest differences observed during moderate activity. Although these differences were attenuated after adjustment for age, the substantial age imbalance between cohorts limits the interpretation of an independent cancer survivorship effect. The activity-dependent pattern nevertheless suggests that differences in autonomic regulation may become more apparent under increased physiological demand. These findings highlight the importance of considering activity context, signal quality, and participant characteristics when interpreting HRV from wearable data in real-world settings.

Several additional limitations should be considered. The imbalance in recording duration between groups may influence statistical comparisons. Treatment-specific stratification was not performed, and potentially relevant factors including BMI, medication use, menopausal status, physical fitness, respiratory rate, posture, speech, and emotional stress were not controlled for in the present analysis. In addition, variability in participant behavior and activity patterns inherent to free-living recordings may affect the consistency of physiological measurements across sessions.

Future work will focus on extending the framework to larger and more diverse populations, incorporating additional physiological and clinical information, and exploring learning-based approaches for joint modeling of signal quality and physiological features.

REFERENCES

- [1] J. F. Thayer, F. Åhs, M. Fredrikson, J. J. Sollers III, and T. D. Wager, "A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health," *Neuroscience & Biobehavioral Reviews*, vol. 36, no. 2, pp. 747–756, 2012.
- [2] X. Zhou, Z. Ma, L. Zhang, S. Zhou, J. Wang, B. Wang, and W. Fu, "Heart rate variability in the prediction of survival in patients with cancer: A systematic review and meta-analysis," *Journal of Psychosomatic Research*, vol. 89, pp. 20–25, 2016.
- [3] J. Zhang, Y. Guo, X. Dong, T. Wang, J. Wang, X. Ma, and H. Wang, "Opportunities and challenges of noise interference suppression algorithms for dynamic ecg signals in wearable devices: A review," *Measurement*, vol. 250, p. 117067, 2025.
- [4] G. D. Clifford, J. Behar, Q. Li, and I. Rezek, "Signal quality indices and data fusion for determining clinical acceptability of electrocardiograms," *Physiological measurement*, vol. 33, no. 9, pp. 1419–1433, 2012.
- [5] U. Satija, B. Ramkumar, and M. S. Manikandan, "A review of signal processing techniques for electrocardiogram signal quality assessment," *IEEE reviews in biomedical engineering*, vol. 11, pp. 36–52, 2018.
- [6] A. Raimondi, A. Busacca, G. C. Giaconia, Y. Antonacci, S. Stivala, L. Faes, and R. Pernice, "Effects of electrocardiographic noise on ultra-short term heart rate variability indices," *Computers in Biology and Medicine*, vol. 196, p. 110803, 2025.
- [7] C. Orphanidou, *Signal quality assessment in physiological monitoring: state of the art and practical considerations*. Springer, 2017.
- [8] F. Marin, K. Lepetit, L. Fradet, C. Hansen, and K. B. Mansour, "Using accelerations of single inertial measurement units to determine the intensity level of light-moderate-vigorous physical activities: Technical and mathematical considerations," *Journal of biomechanics*, vol. 107, p. 109834, 2020.
- [9] P. Gaur, D. S. Temple, M. Hegarty-Craver, M. D. Boyce, J. R. Holt, M. F. Wenger, E. A. Preble, R. P. Eckhoff, M. S. McCombs, H. C. Davis-Wilson *et al.*, "Continuous monitoring of heart rate variability in free-living conditions using wearable sensors: exploratory observational study," *JMIR Formative Research*, vol. 8, p. e53977, 2024.
- [10] R. Cui, X. Zheng, K. Wang, H. Chen, W. Chen, C. Chen, and J. Luo, "Exploring key factors on capacitive ecg signal quality under practical considerations," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–14, 2024.
- [11] J. Rimok and K. A. Larkin-Kaiser, "From lab to life: Wearables, real-world data, and the future of autonomic research," *Autonomic Neuroscience*, vol. 264, p. 103377, 2026.
- [12] V. Rajasekhar, M. Naaz, S. M. R. Naik, B. Nazma, and M. Jyothirmai, "A dl framework to remove noises from ecg signals using cyclic gan," in *2025 IEEE 17th International Conference on Computational Intelligence and Communication Networks (CICN)*. IEEE, 2025, pp. 265–270.
- [13] X. Zhang, J. Li, Z. Cai, L. Zhao, and C. Liu, "Deep learning-based signal quality assessment for wearable ecgs," *IEEE Instrumentation & Measurement Magazine*, vol. 25, no. 5, pp. 41–52, 2022.
- [14] S. Yu, "Research on adaptive denoising method for ecg signal based on k-means algorithm," in *2025 5th International Conference on Artificial Intelligence and Industrial Technology Applications (AIITA)*. IEEE, 2025, pp. 178–181.
- [15] R. Holgado-Cuadrado, C. Plaza-Seco, L. Lovisolio, and M. Blanco-Velasco, "Characterization of noise in long-term ecg monitoring with machine learning based on clinical criteria," *Medical & Biological Engineering & Computing*, vol. 61, no. 9, pp. 2227–2240, 2023.
- [16] M. Peltola, "Role of editing of r-r intervals in the analysis of heart rate variability," *Frontiers in Physiology*, vol. Volume 3 - 2012, 2012.
- [17] G. G. Berntson, J. Thomas Bigger Jr, D. L. Eckberg, P. Grossman, P. G. Kaufmann, M. Malik, H. N. Nagaraja, S. W. Porges, J. P. Saul, P. H. Stone *et al.*, "Heart rate variability: origins, methods, and interpretive caveats," *Psychophysiology*, vol. 34, no. 6, pp. 623–648, 1997.
- [18] W. Wang, M. Shao, W. Du, and Y. Xu, "Impact of exhaustive exercise on autonomic nervous system activity: insights from hrv analysis," *Frontiers in Physiology*, vol. Volume 15 - 2024, 2024.
- [19] E. Caro-Moran, C. Fernandez-Lao, N. Galiano-Castillo, I. Cantarero-Villanueva, M. Arroyo-Morales, and L. Diaz-Rodriguez, "Heart rate variability in breast cancer survivors after the first year of treatments: a case-controlled study," *Biological research for nursing*, vol. 18, no. 1, pp. 43–49, 2016.
- [20] B. L. S. C. Sousa, T. dos Santos Rosa, V. O. Silvino, E. M. L. Barros, H. de Luca Corrêa, D. M. L. do Prado, A. S. Veras-Silva, M. D. de Souza, C. E. B. de Lima, L. F. L. de Oliveira *et al.*, "Physical and autonomic functionality in women with breast cancer pre and post chemotherapy: a case control study," *BMC Sports Science, Medicine and Rehabilitation*, vol. 16, no. 1, p. 5, 2024.
- [21] K. Majerova, M. Zvarik, I. Ricon-Becker, T. Hanalis-Miller, I. Mikolaskova, V. Bella, B. Mravec, and L. Hunakova, "Increased sympathetic modulation in breast cancer survivors determined by measurement of heart rate variability," *Scientific Reports*, vol. 12, no. 1, p. 14666, 2022.
- [22] A. Borsati, L. Toniolo, C. Ciurnelli, F. Bettariga, S. T. Riva, P. Frada, L. Belluomini, I. Trestini, D. Tregnago, J. Insolda *et al.*, "Cardiac autonomic dysfunction in patients with cancer: exploring the role of physical exercise," *Journal of Cancer Research and Clinical Oncology*, vol. 151, no. 6, p. 197, 2025.
- [23] Movesense Oy, "Movesense md." [Online]. Available: <https://www.movesense.com/product/movesense-medical-md/>
- [24] F. Gustafsson, "Determining the initial states in forward-backward filtering," *IEEE Transactions on signal processing*, vol. 44, no. 4, pp. 988–992, 2002.
- [25] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *International conference on pervasive computing*. Springer, 2004, pp. 1–17.
- [26] A. K. Singh and S. Krishnan, "Ecg signal feature extraction trends in methods and applications," *BioMedical Engineering OnLine*, vol. 22, no. 1, p. 22, 2023.
- [27] P. J. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *Journal of computational and applied mathematics*, vol. 20, pp. 53–65, 1987.
- [28] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE transactions on pattern analysis and machine intelligence*, no. 2, pp. 224–227, 2009.
- [29] T. F. o. t. E. S. o. C. t. N. A. S. o. P. Electrophysiology, "Heart rate variability: standards of measurement, physiological interpretation, and clinical use," *Circulation*, vol. 93, no. 5, pp. 1043–1065, 1996.
- [30] J. MacQueen, "Multivariate observations," in *Proceedings of the 5th Berkeley symposium on mathematical statistics and probability*, vol. 1. University of California press Oakland, CA, USA, 1967, pp. 281–297.
- [31] S. Farrokhi, W. Dargie, and C. Poellabauer, "Reliable peak detection and feature extraction for wireless electrocardiograms," *Computers in Biology and Medicine*, vol. 185, p. 109478, 2025.
- [32] E. De Giovanni, T. Teijeiro, G. P. Millet, and D. Atienza, "Adaptive r-peak detection on wearable ecg sensors for high-intensity exercise," *IEEE Transactions on Biomedical Engineering*, vol. 70, no. 3, pp. 941–953, 2022.
- [33] R. J. Ellis, B. Zhu, J. Koenig, J. F. Thayer, and Y. Wang, "A careful look at ecg sampling frequency and r-peak interpolation on short-term measures of heart rate variability," *Physiological measurement*, vol. 36, no. 9, pp. 1827–1852, 2015.
- [34] F. Shaffer and J. P. Ginsberg, "An overview of heart rate variability metrics and norms," *Frontiers in Public Health*, vol. 5, p. 258, 2017.
- [35] B. L. Welch, "The generalization of student's problem when several different population variances are involved," *Biometrika*, vol. 34, no. 1–2, pp. 28–35, 1947.
- [36] J. C. Pinheiro and D. M. Bates, *Mixed-effects models in S and S-PLUS*. Springer, 2000.
- [37] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other," *Annals of Mathematical Statistics*, vol. 18, no. 1, pp. 50–60, 1947.
- [38] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *Journal of the Royal statistical society: series B (Methodological)*, vol. 57, no. 1, pp. 289–300, 1995.
- [39] J. Cohen, *Statistical power analysis for the behavioral sciences*. routledge, 2013.
- [40] J. Zhang, "Effect of age and sex on heart rate variability in healthy subjects," *Journal of manipulative and physiological therapeutics*, vol. 30, no. 5, pp. 374–379, 2007.
- [41] A. Voss, R. Schroeder, A. Heitmann, A. Peters, and S. Perz, "Short-term heart rate variability—influence of gender and age in healthy subjects," *PLoS one*, vol. 10, no. 3, p. e0118308, 2015.